DNA SEQUENCES, PEPTIDES, ANTIBODIES AND VACCINES FOR PREVENTION AND TREATMENT OF SARS

CROSS-REFERENCE TO RELATED PATENT APPLICATIONS

This application claims priority from U.S. Provisional Application 60/480,953, filed 24 June 2003, incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[001] The invention relates to the field of vaccines, therapeutics and antibodies. More particularly, the inventions pertains to DNA, protein, and peptide sequences used to construct vaccines, therapeutics, and antibodies for the prevention and/or treatment of SARS; the vaccines, therapeutics and antibodies created and produced using these novel DNA, protein, or peptide sequences; and a method of treatment and/or prevention of SARS using these vaccines, therapeutics and antibodies for the prevention and/or treatment created and produced using these novel DNA, protein, or peptide sequences.

BACKGROUND OF THE INVENTION

[002] Severe acute respiratory syndrome (SARS) has recently been reported in Asia, North America, and Europe (Rota, PA, Oberste, MS, Monroe, SS et. al. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome 2003 Science 300:1394-1399; Ksiazek, TG, Erdman, D, Goldsmith, CS, et al. A novel coronavirus associated with severe acute respiratory syndrome. 2003 N Engl J Med 348:1953-66; Comments on the reported isolation of viruses related to the SARS coronavirus in wild animals in southern China -available at http://www.who.int/csr/don/2003_05_23b/en/ ; www.cdc.gov/ncidod/sars/ and www.who.int/csr/sars/en/). Data and information on SARS is updated frequently by

the CDC. An update with references of June 4, 2003 by the CDC to the president of the Infectious Disease Society of America was disseminated to the membership of the IDSA and provided most of the background information below.

[003] In March 2003, a novel coronavirus (SARS-CoV) was discovered in association with cases of severe acute respiratory syndrome (SARS). Phylogenetic analyses and sequence comparisons showed that SARS-CoV is not closely related to any of the previously characterized coronaviruses. The WHO has reported that research teams in Hong Kong and Shenzhen, China detected several coronaviruses closely related genetically to the SARS-CoV in two animal species taken from a market (masked palm civet and raccoon-dog). The study also found that one additional species (Chinese ferret badger) elicited antibodies against the SARS coronavirus. Sequencing of viruses isolated from these animals demonstrated that, with the exception of a small additional sequence, the viruses are identical with the human SARS virus. Much more research is needed before any firm conclusions can be reached on the potential role of animals in the transmission of SARS.

[004] During November 1, 2002-May 28, 2003, a total of 8,240 SARS cases were reported to WHO from 28 countries, including the United States; 745 deaths (case-fatality proportion: 9.0%) have been reported. 363 SARS cases have been identified in the United States. There have been reports from 41 states and Puerto Rico, with 297(82%) cases classified as suspect SARS and 66 (18%) classified as probable SARS (more severe illnesses characterized by the presence of pneumonia or acute respiratory distress syndrome) (MMWR 2003;52:500-501). No SARS-related deaths have been reported in the United States. (World Health Organization. Cumulative number of reported cases of severe acute respiratory syndrome (SARS). http://www.who.int/csr/sarscountry/2003_05_28/en.; CDC. Updated interim U.S. case definition of severe acute respiratory syndrome (SARS). http://www.cdc.gov/ncidod/sars/casedefinition.htm; CDC. Severe acute respiratory syndrome (SARS) and coronavirus testing---United States, 2003. 2003 MMWR 52:297-302 http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5214a1.htm; CDC. Update: severe acute respiratory syndrome---United States, 2003. 2003MMWR52:357-60. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5216a1.htm;

[005] Peiris, JSM, Chu, CM, Cheng, VCC et. al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. 2003 Lancet361 1767; CDC. Severe Acute Respiratory Syndrome -- Singapore, 2003. MMWR 2003:52;405-411; CDC. Cluster of Severe Acute Respiratory Syndrome Cases Among Protected Health-Care Workers -- Toronto, Canada, April 2003. 2003 MMWR 52;433-436.

[006] Worldwide, a significant proportion of transmissions have occurred within the healthcare setting, and most transmissions originated from patients for whom infection control precautions had not yet been applied. For most patients the incubation period falls between 4-6 days. To date there has been no evidence of transmission from asymptomatic patients, although transmission may occur early in the course of illness.

[007] Communicability appears to peak in the second week of illness, and viral excretion may peak during this time as well. The duration of communicability following onset of symptoms is unknown. Evidence of SARS-CoV RNA has been detected by RT-PCR in respiratory secretions, urine, and stool for several weeks following symptom onset, although it is unclear whether detection of viral RNA represents the presence of viable virus capable of causing infection. The efficiency of transmission appears to be highly variable between patients. Many cases have been associated with little or no secondary transmission, but a small minority of patients appears to have accounted for many secondary cases. The factors associated with high rates of secondary transmission are poorly understood. In most cases transmission appears to require close contact, most likely via large respiratory droplets or direct contact. Several investigators have reported that SARS-CoV may remain viable on environmental surfaces for 1-2 days, raising the possibility of fomite transmission, but these observations require confirmation. Airborne transmission, although not apparently a major mode of transmission, may have played a role in some reported clusters of infection, particularly during the performance of aerosol generating medical procedures.

[008] *Pathogenesis*: Although the pathogenesis of SARS remains poorly understood, some investigators have hypothesized that the disease process may be, in part, immune mediated, raising the hypothesis that immunomodulatory agents such as corticosteroids may have a role in treatment. No controlled studies have been conducted and the efficacy of this approach is unknown. Some investigators have hypothesized that SARS-CoV-specific immune globulin may be of benefit. However, in some animal coronavirus infections antibody responses may produce an antibody-dependant enhancement of disease.

[009] *Diagnostic Tests*: At this time, tests for SARS-CoV are still being refined, and the sensitivity and specificity are uncertain and still being evaluated. The majority are assays that measure antibodies to the virus. Reverse transcription-polymerase chain reaction (RT-PCR) testing is also available. This test can detect SARS-CoV RNA in clinical specimens, including serum, stool, and nasal secretions. Viral isolation for SARS-CoV also has been done. In these studies, clinical specimens from SARS patients are co-cultured with well-characterized cell lines, and then laboratorians look for evidence of SARS-CoV replication in these cultured cells.

[0010] *Therapy*: All current therapy is supportive. At this time there are no specific therapeutic regimens with proven efficacy. Ribavirin appears to lack in vitro activity against SARS-CoV. Alpha and beta interferons may have some in vitro activity against the SARS-CoV, but the endogenous cytokine response in SARS is poorly understood and it remains unclear whether administration of exogenous interferons might be beneficial. Some 3C proteinase inhibitors have shown promising in vitro activity, but there are no licensed 3C proteinase inhibitors currently available for human use.

[0011] *Prevention*: All current prevention is by identification and isolation of cases. There is no vaccine available. Controlled therapeutic trials are needed.

[0012] *Rationale for the* Invention: A vaccine would be the ideal intervention for preventing SARS. However, there is no vaccine for SARS, and no consensus on the best approach to SARS vaccine development. However, there has already been considerable discussion regarding different approaches to SARS vaccine

development, which was summarized at a May 30, 2003 meeting at the National Institutes of Health (http://www.niaid.nih.gov/sars_meeting.htm).

[0013] It has been demonstrated that antibody interference with virion attachment to target cells is a major neutralization mechanism. (Matsumi, S, et al. Neutralizing monoclonal antibody against an external envelope glycoprotein (gp110) of SIV mac251. 1995 AIDS Res Hum Retroviruses (11) 501-8; Samuelsson, A. et al. Chimeric macaque/human Fab molecules neutralize simian immunodeficiency virus. 1995 Virology (207) 495-502; Sattentau, QJ, et al. Antibody neutralization of HIV-1 and the potential for vaccine design. 1999 Immunol. Lett. (66) 143-9; Wobus, CE, et al. Monoclonal antibodies against the adeno-associated virus type 2 (AAV-2) capsid: epitope mapping and identification of capsid domains involved in AAV-2-cell interaction and neutralization of AAV-2 infection. 2000 J. Virol. (74)9281-93). Evidence supporting a role for the Coronavirus spike protein projections as agents of organ tropism and pathogenesis has been reviewed. (Gallagher, TM and Buchmeier; Coronavirus spike proteins in viral entry and pathogenesis 2001 Virology 279:371-374).

SUMMARY OF THE INVENTION

[0014] Applicants have discovered and created novel gene sequences based on the Spike protein, also called S or E2 glycoprotein, of SARS CoV that, when used to express the polypeptide in a DNA plasmid, recombinant virus, recombinant bacteria, replicon, or other DNA or RNA based vaccine delivery system, or to produce a recombinant protein or synthetic peptide, can enable the creation and production of vaccines, therapeutics, and or antibodies that are useful to treat and/or prevent SARS. In the context of the present invention, a "replicon" is a an autonomously replicating piece of DNA other than the chromosome, such as a plasmid or a phage.

[0015] More particularly, applicants have discovered and created novel gene, novel protein, and novel peptide sequences that can be used to produce vaccines, therapeutics, and/or antibodies that will provide effective treatment or prevention of SARS by their administration on the skin, in the skin, in the subcutaneous tissue, in muscles, in the blood stream, or on any mucosal surface, or by oral ingestion.

[0016] Furthermore, applicants have discovered and created novel gene, protein, and peptide sequences that can be used to produce vaccines, therapeutics, and/or antibodies that will provide effective treatment or prevention of SARS by their administration on the skin, in the skin, in the subcutaneous tissue, in muscles, in the blood stream, or on any mucosal surface, or by oral ingestion, and therefore the method of treating and preventing SARS.

[0017] Accordingly, it is an object of this invention to provide a family of Region I Spike Protein SARS CoV Native no glycosylation (RISPSCoVNNG) polypeptide vaccines that protects a mammal, a human, or an individual or group of individuals against Severe Acute Respiratory Syndrome (SARS).

[0018] It is an object of this invention to provide a family of RISPSCoVNNG polypeptide vaccines that attenuate the effects of SARS in a mammal, a human, or an individual or group of individuals.

[0019] It is an object of this invention to provide DNA vaccines, DNA plasmid vaccines, replicon vaccines, recombinant viral, bacterial and parasitic vaccines encoding the RISPSCoVNNG polypeptide or sub regions thereof that protect a mammal, a human, or an individual or group of individuals against Severe Acute Respiratory Syndrome (SARS).

[0020] It is an object of this invention to provide DNA vaccines replicon vaccines, recombinant viral, bacterial and parasitic vaccines encoding the RISPSCoVNNG polypeptide or sub regions thereof that attenuate the effects of SARS in a mammal, a human, or an individual or group of individuals.

[0021] It is a further object of this invention to provide polypeptides from the Region 1 of the SARS CoV Spike protein in which one or more of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation.

[0022] It is a further object of this invention to provide polypeptides from the Region 1 of the SARS CoV Spike protein in which all potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation.

[0023] It is an additional object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which one or more of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation. .

[0024] It is an additional object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which all of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation.

[0025] It is yet another object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which one or more of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation and, taking advantage of the degeneracy of the DNA code, primary sequence has been altered to optimize expression and enhance immunogenicity for antibodies and thus, protective efficacy.

[0026] It is yet another object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which all of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation and, taking advantage of the degeneracy of the DNA code, the primary sequence has been altered to optimize expression and enhance immunogenicity for antibodies and thus, protective efficacy.

[0027] It is also an object of this invention to provide polyclonal and monoclonal antibodies that recognize, and neutralize polypeptides from the Region I of the SARS CoV Spike protein in which one or more of the potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation.

[0028] It is also an object of this invention to provide polyclonal and monoclonal antibodies that recognize, and neutralize polypeptides from the Region 1 of the SARS CoV Spike protein in which all potential N-linked glycosylation sites have had a peptide residue substituted to prevent glycosylation.

[0029] It is a further object of this invention to provide polypeptides from the Region 1 of the SARS CoV Spike protein in which all potential N-linked glycosylation sites have been changed to Alanine.

[0030] It is an additional object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which one or more of the potential N-linked glycosylation sites have been changed to Alanine.

[0031] It is an additional object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which all of the potential N-linked glycosylation sites have been changed to Alanine.

[0032] It is yet another object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which one or more of the potential N-linked glycosylation sites have been changed to Alanine and, taking advantage of the degeneracy of the DNA code, primary sequence has been altered to optimize expression and enhance immunogenicity for antibodies and thus, protective efficacy.

[0033] It is yet another object of this invention to provide a DNA molecule with a nucleic acid sequence encoding Region I of the SARS CoV Spike protein or sub regions thereof in which all of the potential N-linked glycosylation sites have been changed to Alanine and, taking advantage of the degeneracy of the DNA code, the primary sequence has been altered to optimize expression and enhance immunogenicity for antibodies and thus, protective efficacy.

[0034] It is also an object of this invention to provide polyclonal and monoclonal antibodies that recognize, and neutralize polypeptides from the Region I of the SARS CoV Spike protein in which one or more of the potential N-linked glycosylation sites have been changed to Alanine.

[0035] It is also an object of this invention to provide polyclonal and monoclonal antibodies that recognize, and neutralize polypeptides from the Region 1 of the SARS CoV Spike protein in which all potential N-linked glycosylation sites have been changed to Alanine.

[0036] In consideration of the foregoing objects of the invention and others which will become apparent from the context of the disclosure the Applicants hereby provide the following:

- Novel DNA molecules comprising a nucleotide sequence encoding various sub regions of Region I Spike Protein SARS CoV native no glycosylation (R1SPSCoVNNG).

- Novel Polypeptides of various sub regions of RISPSCoVNNG in which one or more potential N-linked glycosylation sites are substituted with Alanine.

- Vaccines for the prevention, attenuation or treatment of Severe Acute Respiratory Syndrome (SARS) in mammals comprising a DNA molecule that encodes one or more sub regions of R1SPSCoVNNG.

- Vaccines for the prevention, attenuation or treatment of Severe Acute Respiratory Syndrome (SARS) in mammals comprising a sub region of RIIPSCoVNNG

- Antibodies against the novel peptides from RISPSCoVNNG

DESCRIPTION OF THE DRAWINGS

[0037] Figure 1 - Depiction of structure of a typical Coronavirus such as SARS CoV. The structure labeled "S" is the Spike Protein, which is also called the E2 glycoprotein; other features of the viral structure are indicated (H-protein, HE-protein, and RNA genome).

[0038] Figure 2 - Depiction of a host cell interaction with a typical Coronavirus such as SARS CoV.

The virus enters cells via endocytosis and membrane fusion. This critical step in the pathogenesis of the infection and disease is thought to probably be mediated by the E2 glycoprotein, which is represented by the small circles on the surface of the virus.

[0039] Figure 3 - Depiction of the genomic RNA of SARS CoV.

The SARS CoV is a non-segmented, single-stranded, (+) sense RNA, which was 29.736 kb in size in one of the first sequenced isolates. This makes it the longest of any RNA virus currently known. The genome has a 5' methylated cap and 3' poly-A and functions directly as mRNA. The RNA encoding the E2 glycoprotein (also known as the S or Spike protein) is 3.768 kb.

[0040] Figure 4 – Schematic of spike protein fragments (Regions I –IV) that are targeted for expression and DNA plasmid construction, wherein "*" denotes glycosylation sites.

[0041] Figure 5 – Schematic representation of SARS Spike protein *Pichia pastoris* expression constructs depicting portions of the N-terminal and C-terminal sequences of each polypeptide fragment, reiterated in the Table below:

Construct	N-terminus	C-terminus
pPICZαA/SPv5/18	AVDCS	*VFNG*
pPICZαA/SPv8/11	VLY	*GYHL*
pPICZαA/SPv6/33	VLY	*YQDV*
pPICZαA/SP/v7/35	AEQDRN	GYHL

[0042] Figure 6 – SARS expression construct verification by restriction digestion. Ethidium bromide-stained agarose gel shows gene fragments verified by XhoI/XbaI restriction mapping.

[0043] Figure 7 – ELISA titers of rabbits immunized with small peptides from Region I of the native sequence. Rabbits 47299 and 47300 were immunized with 01-11-14-1-pro; rabbits 47297 and 47298 were immunized with 03-11-14-2-pro, as shown in the Table below:

Peptide	Amino Acid Sequence	SEQ ID NO.
03-11-14-1-pro	QILPDPLKPTKRSFI	27
03-11-14-2-pro	TRNIDATSTGNYNYKY	28

[0044] Figure 8 – Image of a Western blot gel showing that polyclonal rabbit sera raised against the native Region I peptides recognized the Spike glycoprotein expressed in mammalian cells: column 1 was loaded with molecular weight standards (Seeblue Plus); column 2, 20 µl of Complete DMEM media+ 10% Fetal Bovine Serum control media; column 3, 20 µl of Complete DMEM media + 10% Fetal Bovine Serum media from HEK293 cell transfected with Spike protein; column 4, 20 Ml of OptiMEM® control media (GIBCO Laboratories, Inc., Grand Island, NY); column 5, 20 µl of OptiMEM® media; column 6, 1.2 µg of Spike protein produced by VRC; column 7, 1.1 µg of Spike protein produced by PP. Sera from each of the rabbits indicated were incubated with the processed gel portions; detection of bound antibodies was carried out using standard Western blot methods.

DESCRIPTION OF THE SEQUENCE LISTING AND TABLES

With regard to sequence numbering the convention followed in this application is as follows: all peptide residue numbering corresponds to the amino acid numbering of the native spike protein sequence – SEQ ID NO: 2; all DNA sequence residue numbering corresponds to the DNA base pair numbering of the native spike protein sequence – SEQ ID NO: 1.

[0045] SEQ ID NO: 1 is a nucleotide sequence encoding SARS Coronavirus Urbani S-protein (corresponding to the gene AY278741).

[0046] SEQ ID NO: 2 is an amino acid sequence for the full length S-protein from accession number AAP13441 (corresponding to the gene AY278741).

[0047] SEQ ID NO: 3 is an amino acid sequence for a fragment of wild-type Spike protein identified herein as Region I (aa 275-1081 of SEQ ID NO: 2).

[0048] SEQ ID NO: 4 is an amino acid sequence for a modified Spike protein fragment corresponding to aa 275-1081 of SEQ ID NO: 2, having 9 glycosylation sites eliminated.

[0049] SEQ ID NO: 5 is a nucleotide sequence encoding Spike protein fragment corresponding to aa 275-1081 of SEQ ID NO: 4, using native virus codons.

[0050] SEQ ID NO: 6 is a nucleotide sequence encoding Spike protein fragment corresponding to aa 275-1081 of SEQ ID NO: 4, using altered codons to enhance expression.

[0051] SEQ ID NO: 7 is the amino acid sequence shown as SEQ ID NO: 2, with potential glycosylation sites indicated.

[0052] SEQ ID NO: 8 is an amino acid sequence of Region II peptide corresponding to amino acids 354-601 as shown in of SEQ ID NO: 4.

[0053] SEQ ID NO: 9 is an amino acid sequence of a peptide corresponding to amino acids 754-1031 as shown in SEQ ID NO: 4 (Region III).

[0054] SEQ ID NO: 10 is an amino acid sequence of a peptide corresponding to amino acids 354-1031 of SEQ ID NO: 4 (Region IV).

[0055] SEQ ID NO: 11 is an amino acid sequence of a peptide corresponding to amino acids 354-601 of SEQ ID NO: 3.

[0056] SEQ ID NO: 12 is an amino acid sequence of a peptide corresponding to amino acids 754-1031 of SEQ ID NO: 3.

[0057] SEQ ID NO: 13 is an amino acid sequence of a peptide corresponding to amino acids 354-1031 of SEQ ID NO: 3.

[0058] SEQ ID NO: 14 is an amino acid sequence of a peptide corresponding to amino acids 275-1031 of SEQ ID NO: 3.

[0059] SEQ ID NO: 15 is a linker for nucleotide primer oligo #1.

[0060] SEQ ID NO: 16 is a linker for nucleotide primer oligo #2.

[0061] SEQ ID NO: 17 is a linker for nucleotide primer oligo #3.

[0062] SEQ ID NO: 18 is a linker for nucleotide primer oligo #4.

[0063] SEQ ID NO: 19 is SARS Spike Glycoprotein Fragment synthetic DNA sequence with N-lined glycosylation site muations (2421 bp).

[0064] SEQ ID NO: 20 is SARS Spike glycoprotein Fragment I protein sequence without the N-linked glycosylation sites (807 amino acids encoded by SEQ ID NO: 19).

[0065] SEQ ID NO: 21 is SARS Spike glycoprotein Fragment II synthetic DNA sequence with N-linked glycosylation site mutations (744bps).

[0066] SEQ ID NO: 22 is SARS Spike glycoprotein Fragment II protein sequence without the N-linked glycosylation sites (248amino acids, encoded by SEQ ID 21).

[0067] SEQ ID NO: 23 is SARS Spike glycoprotein Fragment III synthetic DNA sequence with N-linked glycosylation site mutations (834 bps).

[0068] SEQ ID NO 24 is SARS Spike glycoprotein Fragment III protein sequence without the N-linked glycosylation sites (278 amino acids, encoded by SEQ ID NO: 23).

[0069] SEQ ID NO: 25 is SARS Spike glycoprotein Fragment IV synthetic DNA sequence with N-linked glycosylation site mutations (2034 bps).

[0070] SEQ ID NO: 26 is SARs Spike glycoprotein Fragment IV protein sequence without the N-linked glycosylation sites (678 amino acids, encoded by SEQ ID NO: 25).

DETAILED DESCRIPTION OF THE INVENTION

[0071] This invention pertains to prevention and treatment of the disease caused by the SARS virus. This has been accomplished by the creation and synthesis of unique DNA sequences, unique polypeptides and proteins encoded by these sequences, vaccines and pharmaceutical compositions prepared using these unique DNA and polypeptide sequences, and monoclonal and polyclonal antibodies raised against these unique polypeptides. The strategy in generating these novel sequences was based upon the known sequence of a surface accessible protein referred to as the "spike protein of SARS-CoV, and to sub regions of that protein and polypeptides derived from that protein.

[0072] Furthermore, the applicants have used antibodies derived from these unique sequences to prepare therapeutics and diagnostics by providing and administering a polyclonal or monoclonal antibody against the S protein.. The correlate of protective immunity for virtually all effective vaccines against viruses is the titer of antibodies that neutralizes virus invasion into cells and/or development of the virus within the cells. For antibodies to neutralize virus they in general have to be directed against proteins or glycoproteins that are accessible on the surface of viruses when they are extracellular. The "spike" (also called S or E2) glycoprotein of SARS-CoV is such a target for vaccine development (Fig. 1). This protein is even more attractive as a target because, it is thought to play a critical role in the pathogenesis of the infection and the disease. It is thought to mediate invasion of the virus into host (human) cells.

[0073] Alternative to production and use of the antibodies disclosed herein, the proteins and polypeptides of the present invention are use to stimulate the individual mammal or human to endogenously make antibodies after administration of a vaccine.

THE SPIKE PROTEIN OF SARS

[0074] The spike (S or E2) protein is encoded by 3768 base pairs (SEQ ID NO: 1), which encode a protein of 1256 amino acids (SEQ ID NO: 2). The gene sequence encoding the S protein, also called the Spike Protein and the E2 glycoprotein of SARS Coronavirus Urbani was extracted from the National Center for Biotechnology Information (NCBI) sequence database, accession number AY278741 (SEQ ID NO:

1). This sequence was submitted by Bellini,W.J. et al from the Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton RD, NE, Atlanta, GA 30333, USA.

ANALYSIS

[0075] The applicants analyzed the corresponding protein sequence (NCBI sequence accession number AAP13441)(SEQ ID NO: 2). Based on applicant's experience with identifying biologically active regions of proteins thought to mediate invasion into cells and to be good targets for antibody mediated protective immunity (Sim et al. 1994 Science 264:1941-44; Sim et al. 2001 Molecular Medicine 7:247-254), applicants selected a region of the SARS CoV spike (S or E2) protein to target. This region, which applicants have designated Region I spike protein SARS CoV, is encoded by 2421 base-paired nucleotides (bp 822 to 3243 of SEQ ID NO: 1) and codes for a protein from amino acid 275-1081 (807 amino acids) of the native protein (SEQ ID NO: 3).

[0076] Interestingly, Region I (SEQ ID NO: 3), derived from accession number AAP13441 from CDC was compared with the S-protein sequence from different AAP30030 from Beijing, AAP30713 from Hong Kong, AAP37017 from Taiwan, NP_828851 from Canada. This 275-1081aa region of the S-protein is highly conserved. Among all the sequences published, there is a one amino acid conservative change at amino acid position 577 at which NP_828851 is Alanine and the rest of the sequences are Serine.

POTENTIAL N-LINKED GLYCOSYLATION SITES OF REGION I

[0077] The S-protein sequence contains more than 22 potential N-linked glycosylation sites. According to experiments carried out in support of the present invention, N-glycosylation of such proteins reduces expression of the proteins and reduces immunogenicity for antibodies and thus, protective efficacy. Region I spike protein SARS CoV, the region from amino acid 275 to 1081 (SEQ ID NO: 3), was selected in part because of the paucity of N- glycosylation sites. The region was found to contain two interior domains with no potential N-glycosylation sites. Within the sequences flanking and intervening the two central domains, but included within

amino acids 275-1081, there are 9 potential N-linked glycosylation sites (SEQ ID NO: 7).

[0078] All eukaryotic cells express N-linked glycoproteins. An N-linked gycosylation site requires a consensus sequence Asn-X-Ser/Thr in the primary sequence of the peptide at the putative site . Substitution of any of the amino acid residues in the glycosylation triplet will eliminate that site as a potential N-linked glycosylation site. In the embodiments which follow, alanine has been substituted for residues in the native sequence as indicated to eliminate potential N-linked glycosylation sites, but it will be understood by those skilled in the art that any amino acid substitution of either of the first and third residues in the triplet will accomplish the same goal.

[0079] Based on the Applicant's experience, the polypeptides of native regions I, II, III, IV are considered enhanced for expression of polypeptides and enhanced for reactive immunogenicity for antibody induction and thus, efficacious immunogenic regions of the Spike protein. This results from the infrequency of potential N-linked gycosylation sites identified by the Applicants. The polypeptides of novel regions I, II, III, IV, with potential N-linked glycosylation sites removed, are considered yet more expressible and immunogenic. Therefore, Regions I, II, III, IV polypeptides, with and without the gycosylation sites altered, are used to generate both polyclonal and monoclonal antibodies according to methods known in the art (Shinnick, TM, et al, Peptide-elicited protein-reactive antibodies in molecular biology and medicine (1984) J. Invest. Dermatol. (83) 112s-115s; Sim, KL, et al., Induction of Biologically Active Antibodies in Mice, rabbits, and Monkeys by Plasmodium falciparum EBA-175 Region II DNA vaccine (2001) Molecular Medicine (7) 247-254). These antibodies will be used for both diagnosis and treatment of SARS.

[0080] Inventively, the applicants defined a DNA sequence based on the known sequence of the Region I spike protein SARS CoV and then altered the known sequence to eliminate some or all (from 1 to 9) of these potential N-linked glycosylation sites. The methods for substitution of amino acids is well known in the

art. (Merrifield, R. B. (1963). *J. Am. Chem. Soc.* **85**, 2149–2154. Likewise, methods for producing mutant polynucleotide sequences are known in the art (see, for example, Ausubel et al [1997] Current Protocols in Molecular Biology, John Wiley & Sons, N.Y). In an embodiment, SEQ ID NO: 3 was altered at the amino acid residues corresponding to the following amino acids (as numbered in SEQ ID NO: 2): T^{320} to A, $N^{330}$to A, $T^{359}$to A, $S^{591}$to A, $N^{602}$to A, $N^{691}$to A, $S^{701}$to A, $S^{785}$to A, $T^{1058}$to A to remove the 9 potential N-linked glycosylation sites. This provided a unique polypeptide corresponding to Region I of the spike protein of SARS CoV native polypeptide in which the 9 identified potential N-linked glycosylation sites have been removed by the aforementioned Alanine substitutions. In this context the term "correspond" means that a DNA or polypeptide sequence is the same as a second sequence except for the specified substitutions. This novel polypeptide has been designated "RISPSCoVNNG" (SEQ ID NO: 4). The DNA sequence encoding SEQ ID NO: 4 with native SARS Coronavirus Urbani codon usage was designed and synthesized (R1SPSCoVNNG DNA - SEQ ID NO: 5). Methods of DNA synthesis are also well known in the art (**Uhlmann E. (1988)** Gene. Nov 15;71(1):29-40.

[0081] Applicant's extensive experience with expression of proteins in *Pichia pastoris* indicates that altering the codon usage (the degenerative flexibility of the DNA sequence code means multiple triplets code for the same amino acid) based on applicants' proprietary information leads to enhanced expression of the protein (Narum, DL, et al, Codon Optimization of Gene Fragments Encoding *Plasmodium falciparum* Merzoite Proteins Enhances DNA Vaccine Protein Expression and Immunogenicity in Mice. (2001) Infection and Immunity 69:7250-7253). Furthermore, this alteration of codon usage also enhances in vivo expression of proteins in DNA vaccines. We therefore used the Region I Spike Protein SARS CoV native no glycosylation (SEQ ID NO: 5) as the basis for creating a second synthetic gene in which we altered the sequence to optimize codon usage. This sequence is designated Region I Spike Protein SARS CoV codon optimized no glycosylation "RISPSCoVCONG" (SEQ ID NO: 6). It will be recognized by those skilled in the art that RISPSCoVCNNG and RISPSCoVCONG DNAs both code for the same polypeptide sequence which is designated the RISPSCoVCNNG (Region I) polypeptide (SEQ ID NO: 4).

[0082] These novel synthetic DNA sequences RISPSCoVNNG (SEQ ID NO: 5) and RISPSCoVCONG (SEQ ID NO: 6), and portions and sub regions thereof, are then used to construct novel DNA plasmid vaccines, recombinant virus vaccines, and recombinant bacteria vaccines, and to produce novel recombinant proteins and synthetic peptides that will be used as vaccines. Methods for expressing the polypeptide in a DNA plasmid, recombinant virus, recombinant bacteria, replicon, or other DNA or RNA based vaccine delivery system, or to produce a recombinant protein or synthetic peptide are well known in the art.

[0083] Methods of administration of peptide, DNA or RNA vaccines formed in accordance with the present invention are known in the art. As used herein, the term "administration" or "administering" refers to the process of delivering an agent to a patient, wherein the agent directly or indirectly increases the titer of anti-SARS antibody within the patient, along the lines described in the experimental rabbit model system described in Example 4 herein. The process of administration can be varied, depending on the agent, or agents, and the desired effect. Thus, the process of administration involves administering a selected immunogen of the invention to a patient in need of such treatment. Administration can be accomplished by any means appropriate for the therapeutic agent, for example, by parenteral, mucosal, pulmonary, topical, catheter-based, or oral means of delivery. Parenteral delivery can include for example, subcutaneous intravenous, intrauscular, intra-arterial, and injection into the tissue of an organ. Mucosal delivery can include, for example, intranasal delivery, preferably administered into the airways of a patient, i.e., nose, sinus, throat, lung, for example, as nose drops, by nebulization, vaporization, or other methods known in the art. Oral or intranasal delivery can include the administration of a propellant. Pulmonary delivery can include inhalation of the agent. Catheter-based delivery can include delivery by iontropheretic catheter-based delivery. Oral delivery can include delivery of a coated pill, or administration of a liquid by mouth. Administration can generally also include delivery with a pharmaceutically acceptable carrier, such as, for example, a buffer, a polypeptide, a peptide, a polysaccharide conjugate, a liposome, and/or a lipid, according to methods known in the art.

[0084] Compositions of the subject invention can be formulated according to known methods for preparing pharmaceutically useful compositions. Formulations are described in a number of sources which are well known and readily available to those skilled in the art. For example, Remington's Pharmaceutical Science (Martin EW [1995] Easton Pennsylvania, Mack Publishing Company, 19th Ed.) describes formulations which can be used in connection with the subject invention. Formulations suitable for parenteral administration include, for example, aqueous sterile injection solutions, which may contain antioxidants, buffers, bacteriostats, and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and nonaqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and may be stored in a freeze dried (lyophilized) condition requiring only the condition of the sterile liquid carrier, for example, water for injections, prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powder, granules, tablets, etc. It should be understood that in addition to the ingredients particularly mentioned above, the formulations of the subject invention can include other agents conventional in the art having regard to the type of formulation in question.

[0085] Therapeutically effective and optimal dosage ranges for vaccines and immunogens of the invention can be determined using methods known in the art. Guidance as to appropriate dosages to achieve an anti-viral effect is provided from the exemplified assays disclosed herein. More specifically, results from the immunization pattern described with reference to the rabbit animal model described in Example 4 and Figure 7 herein can be extrapolated by persons having skill in the requisite art to provide a test vaccination schedule. Volunteer subjects or test animals can be inoculated with varying dosages at scheduled intervals and test blood samples can be evaluated for levels of antibody and/or SARS neutralizing activity present in the blood, using the guidance set forth in herein for evaluation of rabbit blood (for example, by Western blot analysis, as depicted in Figure 8). Such results can be used to refine an optimized immunization dosage and schedule for effective immunization of mammalian, specifically human, subjects.

[0086] The polypeptides of native regions I, II, III, IV are considered enhanced immunogenic regions of the Spike protein due to the identified infrequency of potential N-linked glycosylation sites The polypeptides of novel regions I, II, III, IV, with potential N-linked glycosylation sites removed, are considered yet more immunogenic. Therefore, Regions I, II, III, IV polypeptides, with and without the glycosylation sites altered, are used to generated both polyclonal and monoclonal antibodies according to methods known in the art (Shinnick, TM, et al, Peptide-elicited protein-reactive antibodies in molecular biology and medicine (1984) J. Invest. Dermatol. (83) 112s-115s; Sim, KL, et al., Induction of Biologically Active Antibodies in Mice, rabbits, and Monkeys by *Plasmodium falciparum* EBA-175 Region II DNA vaccine (2001) Molecular Medicine (7) 247-254). These antibodies can be used for both diagnosis and treatment of SARS. By way of example, it is contemplated that such antibodies may be used in kits and diagnostic tests for detecting the presence of SARS-CoV in bodily fluids, such as blood, blood fractions, saliva and the like. The anti-SARS antibodies may also be utilized in kits and tests used for environmental surveillance.

[0087] Both the foregoing description and the following Examples are exemplary and explanatory only and are not restrictive of the invention, as claimed. Moreover, the invention is not limited to the particular embodiments described, as such may, of course, vary. Further, the terminology used to describe particular embodiments is not intended to be limiting, since the scope of the present invention will be limited only by its claims.

[0088] With respect to ranges of values, the invention encompasses each intervening value between the upper and lower limits of the range to at least a tenth of the lower limit's unit, unless the context clearly indicates otherwise. Further, the invention encompasses any other stated intervening values. Moreover, the invention also encompasses ranges excluding either or both of the upper and lower limits of the range, unless specifically excluded from the stated range.

[0089] Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of ordinary skill in the art to which this invention belongs. One of ordinary skill in the art will also appreciate that

any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention. Further, all publications mentioned herein are incorporated by reference.

[0090] It must be noted that, as used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents unless the context clearly dictates otherwise. Further, all numbers expressing quantities of ingredients, reaction conditions, % purity, and so forth, used in the specification and claims, is modified by the term "about," unless otherwise indicated. Accordingly, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties of the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits, applying ordinary rounding techniques. Nonetheless, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors from the standard deviation of its experimental measurement.

[0091] Obviously, many modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that, within the scope of the appended claims, the invention many be practiced otherwise than as specifically described.

[0092] The following examples further illustrate the invention. They are merely illustrative of the invention and disclose various beneficial properties of certain embodiments of the invention. These examples should not be construed as limiting the invention.

EXAMPLES

Example 1 - Spike protein (SP) fragment cloning and expression of spike protein in *Pichia pastoris*

SP Region-I gene:

[0093] The 2421 bp gene fragment encoding amino acids 275-1081 (SEQ ID NO: 14) of the Spike protein (NCBI# AAP13441 – SEQ ID NO: 3) with 9 potential N-linked glycosylation site mutations in which alanine was substituted as described above and XhoI and XbaI linker sequences at each end was synthesized. See Figure 4.

SP Region II gene:

[0094] The 744 bp gene fragment encoding amino acids 354—601 (SEQ ID NO: 11) of the NCBI#AAP13441 Spike protein (SEQ ID NO: 3) was amplified by PCR using high fidelity DNA polymerase Vent with primers containing linkers with XhoI and XbaI restriction sites [oligo #1 (SEQ ID NO: 15) and #2 (SEQ ID NO: 16) using SP Region I plasmid as the DNA template (See **Uhlmann E., above**). This is shown in Figure 4

SP Region III gene:

[0095] The 837 bp gene fragment encoding amino acids 754—1031 (SEQ ID NO: 12) of the NCBI#AAP13441 Spike protein (SEQ ID NO: 3) was amplified by PCR using high fidelity DNA polymerase Vent with primers containing linkers with XhoI and XbaI restriction sites [oligo #3 (SEQ ID NO: 17) and oligo #4 (5'- SEQ ID NO: 18)] using SP Region I plasmid as the DNA template. (See **Uhlmann E., above**). This is shown in Figure 4.

SP Region IV gene:

[0096] The 2034 bp gene fragment encoding aa 354-1031 (SEQ ID NO: 13) of the Spike protein (AAP13441; SEQ ID NO: 3) was amplified by PCR using high fidelity DNA polymerase Vent with primers containing linkers with XhoI and XbaI restriction sites [oligo #1 (SEQ ID NO: 15) and oligo #4 (SEQ ID NO: 18)] using SP Region I plasmid as the DNA template. (See **Uhlmann E., above**). This is shown in Figure 4.

[0097] **All 4** gene **fragments** were gel purified, digested with XhoI and XbaI restriction enzymes, ligated into the XhoI/XbaI sites of pPICZaA separately. Applicant's cloning strategy utilized the fact that the target gene is driven by the powerful alcohol oxidase AOX1 promoter that is inducible with methanol. Applicant's cloning strategy also utilizes the alpha factor secretion signal of native

yeast origin, ultimately used for secreting each recombinant Spike protein fragment into the culture media.

[0098] Each ligation mix was transformed into Top10 *E. coli* strain and the plasmid DNAs were analyzed by restriction mapping. Two of the plasmids with the correct gene sequences **for each construct** were linearized and transformed into *P. pastoris* host strain (X33). Three transformants were selected for Zeocin resistance over increasing Zeocin concentrations for each plasmid. Six transformants were picked and screened for their protein expression levels for each construct. Single colonies derived from those transformants were grown in BMGY/BMMY medium (100 mM Potassium phosphate, pH 5.6, 1% yeast extract, 2% peptone, 1.34% yeast nitrogen base without amino acids, $4X10^{-5}$% biotin, 1% glycerol for BMGY and 0.5% methanol for BMMY) and induced at both 24°C and 30°C to test their productivities.

[0099] The supernatants from 24h, 48h, 72h, 96h and 120h post-induction time points were analyzed on both Coomassie stained SDS-PAGE gels and immunoblots (using sera generated from immunization of synthetic peptides in rabbits).

Example 2 - Preparation of the P. pastoris clone glycerol cell bank stock

[00100] A single colony from the plate with Zeocin plate will be inoculated into BMGY and grown overnight at 30 ° C at 250 rpm. Sterility of the culture will be examined by light microscopy (400x magnification). A glycerol stock (P1) will be prepared in 15% glycerol and stored at –70°C. A P2 glycerol stock will be prepared by inoculating 100 µl of P1 stock into 100 ml of BMGY and grown overnight at 30°C at 250 rpm and saved in a final of 15% glycerol and stored at –70°C.

Example 3 - Recombinant protein expression by shake-flask fermentation

[00101] One vial of *P. pastoris* clone P2 glycerol stock is inoculated into 100 ml of BMGY in a 500 ml flask and incubated at 250 rpm at 30°C for ~24 h to OD_{600} 5-10. 15ml of this culture is inoculated into 1L BMGY in a 4L-baffled flask and grown at

the same condition for another 24h. The culture is centrifuged at 5000 rpm in a GSA rotor for 5 min and the cell pellets is resuspended in BMMY at 80 g/L cell density. The concentrated cell culture is placed into a 4L-baffled flask and induced at the selected temperature with 250 rpm agitation. The culture is fed with 2% methanol every 24h and the supernatant is collected at the selected post-induction time point by centrifugation at 5000 rpm for 5 minutes, and stored at -70°C.

Example 4 - T cell epitope prediction:

[00102] Using a T cell epitope prediction algorithm applicants have predicted 9 HLA A0201, A0301 or B0801 (class I) and 6 DRB1 (class II) T-cell epitopes in this region of the SARS CoV spike protein. The sequences of thirteen of the 15 epitopes were completely conserved in the sequences of 13 isolates of SARS CoV that we were able to compare. These data indicate that, 1) There are adequate T-cell helper epitopes in the protein that we are working on to insure a broad antibody response in immunized humans, and 2) There are adequate CD8 T-cell epitopes to insure that at least 90% of the world's population will be able to mount CD8 T-cell responses to this portion of the protein. Such data will provide the foundation for parallel T-cell studies as we move forward with vaccine development.

[00103] Applicants selected potential B-cell epitopes and synthesized peptides from native SEQ ID NO: 3. These peptides were conjugated to Keyhole limpet hemocyanin and used to immunize rabbits with Freund's adjuvant. The rabbits were bled 2 weeks after the third dose for sera containing polyclonal antibodies against the peptides. The titers found in the rabbits against the peptides are shown in Figure 4. Immunoblots (Westerns) show that the polyclonal sera are specific against the Spike protein (Figure 5).

[00104] In the foregoing, the present invention has been described with reference to suitable embodiments, but these embodiments are only for purposes of understanding the invention and various alterations or modifications are possible so long as the present invention does not deviate from the claims that follow.

[00105] All references cited, as well as the provisional application upon which this application is based, are incorporated herein by reference. All nucleotide and amino acid sequences mentioned herein or appended hereto are incorporated herein by reference.